

Spreading Signals in the Wideband Limit

Elchanan Zwecher and Dana Porrat
The Hebrew University of Jerusalem
Jerusalem, Israel 91904
elchanzw, dporrat@cs.huji.ac.il

Abstract—Wideband communications are impossible with signals that are spread over a very large band and are transmitted over multipath channels unknown ahead of time. This work exploits the I-mmse connection to bound the achievable data-rate of spreading signals in wideband settings, and to conclude that the achievable data-rate diminishes as the bandwidth increases due to channel uncertainty. The result applies to all spreading modulations, i.e. signals that are evenly spread over the bandwidth available to the communication system, with SNR smaller than $\log(W/L)/(W/L)$ and holds for communications over channels where the number of paths L is unbounded by sub-linear in the bandwidth W .

I. INTRODUCTION

This work analyzes the performance of wideband communication systems, we characterize the SNR regime that does not allow communications in the wideband limit. The underlying reason for the inability to communicate is channel uncertainty, and our result applies to channels where the number of multipath components is unbounded and sub-linear in the bandwidth. Our proof is based on the I-mmse connection [3] and hinges on a calculation of the minimum mean square error (mmse) estimate of the unknown channel.

We consider signals that spread their power over the entire bandwidth, such as PPM or impulse radio [14], where the pulse shape and duration determine the bandwidth, or direct sequence spread spectrum, where the chip duration determines the frequency spread of the signal. Our result applies also to OFDM-type signals, if the entire available bandwidth is used concurrently. Examples of signals that are not spread over bandwidth are FSK and multi-tone FSK [6], [5], where each symbol concentrates power on a small span of frequencies, although the entire range of symbols may span a very large bandwidth.

Our result shows that channel uncertainty is detrimental to spreading systems operating over multipath channels where the number of apparent paths is unbounded but sub-linear with the system bandwidth [11], [10], essentially because of the signal uses too many eigen-modes of the channel. Our model of channel variation in time is a simplistic block coherent one, where the channel is fixed for known lengths of time (coherence periods) and realizations over different coherence periods are IID. This channel model offers the advantage of eigen-modes that are particularly simple, as harmonic signals are eigen-modes of any linear time invariant channel. The channel uncertainty a communication system faces when operating over the block-coherent channel is thus limited to the eigen-values of the channel, or in other words to

the complex channel gain over the frequency band the system uses.

The essential feature of spreading signals that renders them ineffective over wide bands, is that they use the entire range of channel eigen-modes concurrently, and are thus exposed to uncertainty of a large number of parameters (channel eigen-values). Modulation schemes of the FSK type, that exploit a small number of channel eigen-modes per symbol, are exposed to uncertainty in only a small number of parameters.

Our result can be extended to more complex channels, where the variation in time is described using the Doppler spectrum rather than by block-coherence [9], [12], [2]. The eigen-modes of such channels are approximately given by orthogonal Weyl-Heisenberg bases [4] in the under-spread case, i.e. when the channel's response is highly concentrated in the delay-Doppler plain. The essential feature that determines whether communications are possible in the wideband limit is the spreading of symbol power over the unknown channel eigen-modes. In channels where the eigen-modes are not known in advance, it is very difficult to overcome channel uncertainty.

Related work has shown that pulse-position modulation (PPM) systems are unable to handle uncertainty in the delays of multipath components [7], [8]. We showed that the receiver is unable to detect the channel paths if the bandwidth is large enough, whether it uses a threshold detector [7] or a maximum likelihood detector [8]. This work extends the scope of past results to a wider family of signals, and makes a statement on the achievable rate.

The comparison of our result to that of *Telatar & Tse* [13] is also interesting. The *T&T* results discussed continuous signals, i.e. effectively using a signal to noise ratio (SNR) that inversely depends on the system bandwidth $\text{SNR} = \theta(\frac{1}{W})$, whereas our result is more general in the sense that it specifies the range of SNR dependencies on bandwidth that does not allow communications in the wideband regime. In other words, our result applies to impulsive systems and implicitly indicates the minimal level of impulsiveness that allows communications in the limit of large bandwidth.

Another point of comparison of our result to that of [13] is the type of signals to which they apply. *T&T* consider two types of spreading signals: IID complex signals and signals with a very low cross-correlation: $\sum_i X_i X_{i-n} \sim \text{const}$ that does not depend on the length of the vector. Our result applies to a wider family of signals, where the cross correlation may

as high as $\theta(\sqrt{W})$. This is significant because spreading signals generated by IID or pseudo-random sequences have an empirical correlation that varies as $\theta(\sqrt{W})$.

The results in [8], that connect the number of channel paths to the level of impulsiveness are also relevant here. The impulsiveness parameter there is easily translated to the SNR per active burst of transmission, and the results in [8] basically determine a lower bound on the SNR, above which the channel uncertainty penalty is insignificant in the wideband limit. Our new result involves an upper bound on the SNR, below which the channel uncertainty penalty prevents communications.

II. MODEL

We consider communication systems with a (single sided) bandwidth W , operating over block-coherent multipath channels. We use a real discrete model of the system, after sampling at the receiver at rate W . The model over a single coherence period is given by

$$Y = \sqrt{\text{SNR}} X \star \tilde{H} + Z \quad (1)$$

where Y is the received signal over an entire coherence period of length T_c , this is a vector with $K_c = T_c W$ entries. The vector X of length K_c and average energy K_c represents the transmitted signal, the multipath channel is represented by the vector \tilde{H} of length K_c and \star marks a convolution. Z is white standard Gaussian noise (IID with zero mean and variance one) and SNR is the signal to noise ratio. The SNR can be understood as the signal to noise ratio per frequency resolution bin or per degree of freedom. We neglect in (1) the edge effects at the beginning of the coherence period.

We impose a probabilistic energy constraint:

$$P(\|X\|^2 > (1 + o(1)) K_c) \xrightarrow{W \rightarrow \infty} 0 \quad (2)$$

Note that IID signaling satisfies this assumption.

The transmitted signal may be impulsive, we consider the signal to noise ratio during active transmission so there is no need to explicitly address the impulsiveness used by the system (i.e. the duty cycle ratio). We assume that the transmitter does not use information on the channel realization, and the transmitted signal does not depend on it.

The transmitted signal X is wideband: its empirical auto-correlation is upper bounded by

$$\left| \langle \overline{X^i}, \overline{X^j} \rangle \right| \leq B_4 \sqrt{K_c} \quad i \neq j \quad i, j = 1, \dots, K_c \quad (3)$$

with B_4 is a constant that does not depend on the bandwidth. The notation $\langle \cdot, \cdot \rangle$ is used for the inner product of vectors, and the notation $\overline{X^i}$ is used for a vector X that is cyclicly shifted by i positions, i.e.

$$\overline{X^i} = \begin{pmatrix} X_{(1-i)} \\ X_{(2-i)} \\ \vdots \\ X_{(K_c-i)} \end{pmatrix} \quad (4)$$

where $(-)$ indicates a mod K_c difference.

The channel is composed of L paths, each with a delay in the range $[0, T_d W]$ where T_d is the delay spread. The channel is block-constant with coherence time T_c , i.e. it has IID realizations over different coherence periods. We assume $T_d \ll T_c$ and thus justify to an extent our loose treatment of edge effects at the beginning of each coherence period. We approximate (1) with a circularly-shifted matrix:

$$Y = \sqrt{\text{SNR}} x \tilde{H} + Z \quad (5)$$

where

$$x = \begin{pmatrix} X_1 & X_{K_c} & \dots & X_2 \\ X_2 & X_1 & X_{K_c} & \dots \\ \vdots & & \ddots & \ddots \\ X_{K_c} & \dots & X_2 & X_1 \end{pmatrix} = \begin{pmatrix} \overline{X^0} & \overline{X^1} & \dots & \overline{X^{K_c-1}} \end{pmatrix} \quad (6)$$

The channel model is real, L channel gains are IID and zero mean, with variance $1/L$, so the energy in the channel's impulse response equals one on average. We assume an upper bound on path gains $|H_i| > \frac{B_1}{\sqrt{L}}$, with a constant B_1 that does not depend on the bandwidth. The choice of the L non zero taps is uniform over the $\binom{K_c}{L}$ possibilities.

The number of paths L diverges as the bandwidth increases in a sub-linear manner [11], [10], i.e. $L \xrightarrow{W \rightarrow \infty} \infty$ and $L/W \xrightarrow{W \rightarrow \infty} 0$.

We make a probabilistic assumption on the channel's response

$$P\left(\left|\sum_{j=1, j \neq i}^{K_c} \tilde{H}_j \langle \overline{X^i}, \overline{X^j} \rangle\right| > B_3 \sqrt{K_c}\right) \xrightarrow{W \rightarrow \infty} 0 \quad i = 1, 2, \dots, K_c \quad (7)$$

with a constant B_3 that does not depend on the bandwidth. The typical value of the correlation in (7) is $\sqrt{K_c}$, so this assumption is a natural one. By taking a large constant B_3 we ensure that our result holds for almost all values of i .

III. RESULT

Theorem 1: Communication systems modeled by (5), that use spreading signals and operate over multipath channels as described in Section II, with $\text{SNR} \ll \frac{\log W/L}{W/L}$ have a diminishing rate in the limit of large bandwidth:

$$\frac{I(Y; x)}{\frac{1}{2} K_c \text{SNR}} \xrightarrow{W \rightarrow \infty} 0$$

with probability 1.

We prove the theorem in Section V by showing that

$$\lim_{W \rightarrow \infty} \frac{I(Y; \tilde{H}|x)}{\frac{1}{2} K_c \text{SNR}} = 1$$

and applying

$$I(Y; x) = I(Y; \tilde{H}, x) - I(Y; \tilde{H}|x) \quad (8)$$

$$\leq \frac{1}{2} K_c \text{SNR} - I(Y; \tilde{H}|x) \quad (9)$$

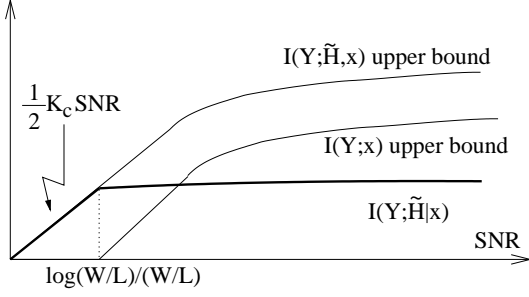


Fig. 1. A sketch of the I-SNR relationship for spread signals for a very large bandwidth. The coherent datarate upper bound (top graph) is linear in the low SNR regime, and convex. The channel uncertainty penalty in the bottom graph is linear for low SNR values, and saturates at the channel's entropy. The incoherent datarate (middle graph) is not convex.

The term $I(Y; \tilde{H}|x)$ in (9) is the datarate penalty due to channel uncertainty.

IV. DISCUSSION

The proof of Theorem 1 is based on calculating the mmse estimate of the channel response \tilde{H} , given the transmitted and the received signals. We show that this mmse estimate is a vector with an $o(1)$ norm in low SNR conditions, essentially because the noise Z overwhelms the information carrying signal.

Theorem 1 shows that in the wideband limit, the low SNR regime can be divided in to parts:

$$\text{SNR} \ll \frac{\log K_c/L}{K_c/L} \quad (10)$$

where spreading signals are not effective, and

$$\text{SNR} > \frac{\log K_c/L}{K_c/L} \quad (11)$$

where although the SNR diminishes in the limit, it enables a positive datarate.

The channel uncertainty penalty is upper bounded by the channel entropy, and the bottom graph of Figure 1 thus saturates at $\text{SNR} = \frac{\log K_c/L}{K_c/L} + \frac{o(L)}{K_c}$, where the first part corresponds to the entropy of the paths' delays and the second to their gains.

V. PROOF OF THEOREM 1

The proof is based on the I-mmse connection, in particular Theorem 2 of [3]. This theorem gives a simple formula to the achievable rate of communications over a known vector channel in terms of the error of the mmse estimate of the transmitted signal. We reverse the roles of H and x in our usage of Theorem 2 of [3], i.e. consider x as known and H as the estimated party. Using our notation, the I-mmse connection says that as long as the vector H satisfies $E\|H\|^2 < \infty$ we have

$$I(H; \sqrt{\text{SNR}} xH + Z) = \frac{1}{2} \int_0^{\text{SNR}} \text{mmse}(\text{SNR}) d\text{SNR} \quad (12)$$

where $\text{mmse}(\text{SNR})$ is given by

$$\text{mmse}(\text{SNR}) = E \left[\|x\tilde{H} - x\hat{H}(Y; \text{SNR})\|^2 \right] \quad (13)$$

and \hat{H} is the mmse estimate of \tilde{H} given both x and Y . We will show that the mmse estimate is a vector with an $o(1)$ norm in low SNR conditions, thus the minimal mean square error (13) converges to K_c , the mutual information (12) converges to $\frac{1}{2} K_c \text{SNR}$ and (9) diminishes.

The mmse estimate is given by

$$\hat{H} = E[H|Y, x] \quad (14)$$

We lower bound the mmse by calculating the minimum mean square error in a system that is given additional information on \tilde{H} , namely which of its positions satisfies (7).

$$\hat{H}' = E[H|Y, x, I_{(7)}] \quad (15)$$

where $I_{(7)}$ is a list of indices $\{i\}$ where \tilde{H}_i satisfies (7). The additional information can only reduce the mmse.

$$\hat{H}' = \int \int \dots \int H f(H|Y, x) dH_1 dH_2 \dots dH_L \quad (16)$$

The conditional probability density in (16) is manipulated using the independence of the transmitted signal from the channel.

$$\hat{H}' = \frac{\int \int \dots \int H f(Y|x, H) f(H) dH_1 dH_2 \dots dH_L}{\int \int \dots \int f(Y|x, H) f(H) dH_1 dH_2 \dots dH_L} \quad (17)$$

The conditional probability density in (17) is Gaussian as $Y|x, H \sim \mathcal{N}(xH, I)$. We denote by $f_s(\cdot)$ the probability density of a K_c long vector of IID standard Gaussian variables

$$f_s(S) = \frac{1}{(2\pi)^{K_c/2}} \exp\left(-\frac{1}{2} \|S\|^2\right) \quad (18)$$

and proceed to examine the components of the vector \hat{H}' . Consider first positions (indices) j where $\tilde{H}_j = 0$. At these positions, any non-zero value of \hat{H}'_j increases the estimation error and can be disregarded in the calculation of a lower bound on the mmse. Let us examine the positions where $\tilde{H}_i \neq 0$ and (7) holds, and look at the estimates of each such value:

$$\begin{aligned} \hat{H}'_i &= \int \int \dots \int H_i f_s\left(\sqrt{\text{SNR}} x (\tilde{H} - H) + Z\right) \\ &\quad f(H) dH_1 dH_2 \dots dH_L / \\ &\quad \int \int \dots \int f_s\left(\sqrt{\text{SNR}} x (\tilde{H} - H) + Z\right) \\ &\quad f(H) dH_1 dH_2 \dots dH_L \end{aligned} \quad (19)$$

We prove that the mmse estimate (19) is very small by comparing the integral in the nominator to the integral in the denominator, that effectively sums over a bigger group of assignments of H . Assuming (7) we show that the nominator of (19) is negligible when compared to its denominator.

We first approximate both integrals in (19) by sums over sampled groups of values of each positions in the vectors H .

The sampling is done over a tight enough grid that the resulting errors are small.

An upper bound to (19) is calculated by breaking the sum in the denominator to a series of sums over groups of values of H , where each group corresponds to a single assignment of H in the nominator. The set \mathcal{B} consists of assignments with a non-zero value $H_i \neq 0$. Rewriting the discrete approximation of (19) we get

$$\hat{H}'_i = \frac{\sum_{H \in \mathcal{B}} H_i f_s \left(\sqrt{\text{SNR}} x (\tilde{H} - H) + Z \right) p(H)}{\sum_{G \in \mathcal{B} + \mathcal{B}^c} f_s \left(\sqrt{\text{SNR}} x (\tilde{H} - G) + Z \right) p(G)} \quad (20)$$

the notation G was introduced to improve clarity we define $p(H) = f(H)\Delta H^L$ and ΔH is the sampling step of each dimension of the vector H .

We proceed to divide the entire range of vectors G into non-overlapping subgroups, such that for each $H \in \mathcal{B}$ we have a corresponding subgroup $\mathcal{A}(H)$, such that

$$J(H) = \frac{H_i f_s \left(x (\tilde{H} - H) + Z \right) p(H)}{\sum_{G \in \mathcal{A}(H)} f_s \left(x (\tilde{H} - G) + Z \right) p(G)} \quad (21)$$

diminishes in the limit of large bandwidth. The convergence of (20) to zero follows directly, as the nominator of (20) is a sum of the nominators of $J(H)$ for all $H \in \mathcal{B}$.

The subgroups $\mathcal{A}(H)$ are created randomly. For each assignment of G that has a non-zero value in the i^{th} position, it is put in the subgroup $\mathcal{A}(G)$. For a vector G with $G_i = 0$ we (uniformly) choose one of its non-zero taps and replace it to the i^{th} position. To clarify the process, let us say that the j^{th} position of the vector G was chosen. We calculate a new vector H by

$$H_i = G_j ; \quad H_j = 0 ; \quad H_k = G_k \text{ for } k \neq i, j \quad (22)$$

and assign the vector G to the subgroup $\mathcal{A}(H)$.

Each group $\mathcal{A}(H)$ contains H and about $(K_c - L)/L$ other members, each different from H in exactly two positions. We ensure that groups' sizes do not deviate significantly from $(K_c - L)/L$ by relocating members from large groups into suitable smaller ones.

We denote by $H^{i \rightarrow k}$ a member of $\mathcal{A}(H)$ that differs from H by exchanging the values in its i^{th} and k^{th} positions, and define $H^{i \rightarrow i} = H$. The set $\mathcal{K}(H)$ holds the values of k such that $H^{i \rightarrow k} \in \mathcal{A}(H)$.

The terms $p(H)$ in the nominator of (21) and $p(G)$ in the denominator are identical for all members of the group $\mathcal{A}(H)$ because of our assumptions on IID gains and a uniform spread of the path delays.

$$J(H) = \frac{H_i \exp \left\{ -\frac{1}{2} \left\| Y - \sqrt{\text{SNR}} x H^{i \rightarrow i} \right\|^2 \right\}}{\sum_{k \in \mathcal{K}(H)} \exp \left\{ -\frac{1}{2} \left\| Y - \sqrt{\text{SNR}} x H^{i \rightarrow k} \right\|^2 \right\}} \quad (23)$$

The denominator of (23) contains a sum over about K_c/L exponents with different values of k , including $k = i$ and

the nominator holds a signal such factor with $k = i$. We take a close look at their exponent and introduce the notation $I(H_i, k)$ for a K_c -long vector with the value H_i at the k^{th} positions and zeros elsewhere.

$$- \frac{1}{2} \left\| Y - \sqrt{\text{SNR}} x H^{i \rightarrow k} \right\|^2 = - \frac{1}{2} \|Y\|^2 \quad (24)$$

$$- \frac{1}{2} \left\| \sqrt{\text{SNR}} x (H^{i \rightarrow k} - I(H_i, k)) \right\|^2 \quad (25)$$

$$- \frac{1}{2} \left\| \sqrt{\text{SNR}} x I(H_i, k) \right\|^2 \quad (26)$$

$$- \left\langle \sqrt{\text{SNR}} x (H^{i \rightarrow k} - I(H_i, k)), \sqrt{\text{SNR}} x I(H_i, k) \right\rangle \quad (27)$$

$$+ \left\langle \sqrt{\text{SNR}} x \tilde{H}, \sqrt{\text{SNR}} x (H^{i \rightarrow k} - I(H_i, k)) \right\rangle \quad (28)$$

$$+ \left\langle \sqrt{\text{SNR}} x \tilde{H}, \sqrt{\text{SNR}} x I(H_i, k) \right\rangle \quad (29)$$

$$+ \left\langle Z, \sqrt{\text{SNR}} x (H^{i \rightarrow k} - I(H_i, k)) \right\rangle \quad (30)$$

$$+ \left\langle Z, \sqrt{\text{SNR}} x I(H_i, k) \right\rangle \quad (31)$$

we now deal with each line (24)-(31) separately, to show that the nominator of (23) is much smaller than the denominator.

(24): The term $-\frac{1}{2} \|Y\|^2$ does not depend on k .

(25), (28), (30): These terms do not depend on k because the vectors $H^{i \rightarrow k} - I(H_i, k)$ are identical over $k \in \mathcal{K}(H)$.

(26): The term $-\frac{1}{2} \left\| \sqrt{\text{SNR}} x I(H_i, k) \right\|^2$ depends on k , but the norm is constant over $k \in \mathcal{K}(H)$. The vector $I(H_i, k)$ essentially extracts a single column of the matrix x and multiplies it by H_i . The matrix x is circularly symmetric and thus (26) is fixed.

(27): The term

$$\begin{aligned} a_k &= - \left\langle \sqrt{\text{SNR}} x (H^{i \rightarrow k} - I(H_i, k)), \sqrt{\text{SNR}} x I(H_i, k) \right\rangle \\ &= -\text{SNR} H_i \sum_{j=1, j \neq k}^{K_c} H_j \left\langle \overline{X^{j-1}}, \overline{X^{k-1}} \right\rangle \end{aligned}$$

is significantly smaller than (29) at the nominator of (23), or in other words an order of magnitude smaller than $\frac{K_c \text{SNR}}{L}$. We prove this by calculating

$$a_i = -\text{SNR} H_i \sum_{j=1, j \neq i}^{K_c} H_j \left\langle \overline{X^{j-1}}, \overline{X^{i-1}} \right\rangle \quad (32)$$

The typical value is on the order of $B_1 \text{SNR} \frac{\sqrt{K_c}}{\sqrt{L}}$ and a loose upper bound is given by Using condition (7) on \tilde{H} , that basically ensures a low correlation in (32), we have

$$|a_i| \leq B_1 B_3 \text{SNR} \sqrt{\frac{K_c}{L}} \quad (33)$$

(29): The term

$$\begin{aligned} b_k &= \left\langle \sqrt{\text{SNR}} x \tilde{H}, \sqrt{\text{SNR}} x I(H_i, k) \right\rangle \\ &= \text{SNR} H_i \sum_{j=1}^{K_c} \tilde{H}_j \left\langle \overline{X^{j-1}}, \overline{X^{k-1}} \right\rangle \end{aligned}$$

is the dominant term in the nominator of (23), i.e. for $k = i$.

$$b_i = \text{SNR} H_i \sum_{j=1}^{K_c} \tilde{H}_j \left\langle \overline{X^{j-1}}, \overline{X^{i-1}} \right\rangle \quad (34)$$

The dominant term in the sum is the i^{th} , where $\left\langle \overline{X^{i-1}}, \overline{X^{i-1}} \right\rangle \leq (1 + o(1)) K_c$ with high probability.

$$\begin{aligned} |b_i| &\leq B_1^2 (1 + o(1)) \text{SNR} \frac{K_c}{L} + B_1 B_3 \text{SNR} \frac{\sqrt{K_c}}{\sqrt{L}} \\ &= B_1^2 (1 + o(1)) \text{SNR} \frac{K_c}{L} \end{aligned} \quad (35)$$

the last approximate equality is tight in the limit of large bandwidth.

(31): The term $c_k = \left\langle Z, \sqrt{\text{SNR}} x I(H_i, k) \right\rangle$ is the dominant term in the denominator of (23). The sum of exponents of $\{c_k\}$ is lower bounded by a single exponent with $k^* \in \mathcal{K}(H)$. We use asymptotic order statistics to show that there is $k^* \in \mathcal{K}(H)$ such that $c_{k^*} = \sqrt{\frac{K_c \text{SNR}}{L}} \sqrt{2 \log \frac{K_c}{L}}$ in the limit. To prove the existence of k^* we examine the joint probability density of $\{c_k\}$. These are mutually Gaussian zero mean random variables, with variance

$$\text{var}(c_k) = H_i^2 \text{SNR} \left\langle \overline{X^k}, \overline{X^k} \right\rangle \leq H_i^2 \text{SNR} (1 + o(1)) K_c$$

and covariance

$$\text{cov}(c_k, c_m) = H_i^2 \text{SNR} \left\langle \overline{X^k}, \overline{X^m} \right\rangle \leq H_i^2 \text{SNR} B_4 \sqrt{K_c}$$

We collect $\{c_k\}$ into the vector C of length $M = |\mathcal{K}(H)|$ and mark its correlation matrix by R_c . R_c is positive definite, it has a constant and large value on its diagonal, and significantly smaller values off-diagonal.

The mean and variance of the maximal of M IID $\sim N(0, \sigma^2)$ random variables are given by [1]: the mean equals

$$\sigma \left(\sqrt{2 \ln M} - \frac{\ln \ln M + \ln 2\pi - 2\mathcal{C}}{2\sqrt{2 \ln M}} + O\left(\frac{1}{\ln M}\right) \right) \quad (36)$$

and the variance is

$$\frac{\pi^2 \sigma^2}{12 \ln M} + O\left(\frac{1}{\ln^2 M}\right) \quad (37)$$

where $\mathcal{C} \approx 0.5772$ is Euler's constant. Note that for a large M the variance diminishes. These results cannot be directly applied to the maximal $\{c_k\}$ because these variables are correlated. We show that the correlations among $\{c_k\}$ are insignificant in the limit of large bandwidth in the sense that there is a c_{k^*} that is very similar to the maximal of IID Gaussians, and conclude that in the limit

$$c_{k^*} \rightarrow (1 + o(1)) \sqrt{H_i^2 \text{SNR} K_c} \sqrt{2 \ln M} \quad (38)$$

In the nominator of (23) the term c_i is insignificant, it has zero mean and a small variance on the order of $\sqrt{\frac{K_c \text{SNR}}{L}}$.

To summarize the discussion of (24)-(31), we can upper bound (23) in the limit of large bandwidth using the significant terms in the nominator and denominator:

$$\begin{aligned} J(H) &\leq \frac{H_i \exp \left\{ -\frac{1}{2} \left\| Y - \sqrt{\text{SNR}} x H^{i \rightarrow i} \right\|^2 \right\}}{\max_{k \in \mathcal{K}(H)} \exp \left\{ -\frac{1}{2} \left\| Y - \sqrt{\text{SNR}} x H^{i \rightarrow k} \right\|^2 \right\}} \\ &\approx H_i \exp \{a_i + b_i - c_{k^*}\} \end{aligned} \quad (39)$$

$$\leq H_i \exp \left\{ 3B_1^2 \frac{K_c \text{SNR}}{L} - \sqrt{\frac{K_c \text{SNR}}{L}} \sqrt{2 \log \frac{K_c}{L}} \right\} \quad (40)$$

and for $\text{SNR} = o\left(\frac{\log \frac{K_c}{L}}{\frac{K_c}{L}}\right)$ the exponent of (40) diverges to $-\infty$ as the bandwidth increases, and $J(H) \xrightarrow{W \rightarrow \infty} 0$. Replacing K_c by WT_C , the proof of Theorem 1 is complete.

REFERENCES

- [1] Harald Cramér. *Mathematical Methods of Statistics*. Princeton University Press, 1946.
- [2] Giuseppe Durisi, Helmut Bolcskei, and Shlomo Shamai. Capacity of underspread WSSUS fading channels in the wideband regime. In *IEEE International Symposium on Information Theory*, pages 1500–1504, Jul. 2006.
- [3] Dongning Guo, Shlomo Shamai, and Sergio Verdú. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Transactions on Information Theory*, 51(4):1261–1282, Apr. 2005.
- [4] Werner Kozek. On the transfer function calculus for underspread LTV channels. *IEEE Transactions on Signal Processing*, 45(1):219–223, Jan. 1997.
- [5] C. Luo, M. Médard, and L. Zheng. Error exponents for multitone frequency shift keying on wideband rayleigh fading channels. In *IEEE Globecom*, volume 2, pages 779–783, Dec. 2003.
- [6] Cheng Luo and Muriel Médard. Frequency-shift keying for ultrawideband - achieving rates of the order of capacity. In *40th Annual Allerton Conference on Communication, Control, and Computing*, 2002.
- [7] Dana Porrat and Urbashi Mitra. Timing acquisition of wideband PPM systems over multipath. In *IEEE International Symposium on Information Theory*, pages 2583–2587, Dec. 2006.
- [8] Dana Porrat, David Tse, and Şerban Nacu. Channel uncertainty in ultra wideband communication systems. *IEEE Transactions on Information Theory*, 53(1):194–108, Jan. 2007.
- [9] V. Raghavan, G. Hariharan, and A. M. Sayeed. Capacity of sparse multipath channels in the ultra-wideband regime. *Selected Topics in Signal Processing*, 1(3):357–371, Oct. 2007.
- [10] F. Saadane, T. Aboutajdine, A.M. Hayar, and R. Knopp. On the estimation of the degrees of freedom of indoor UWB channel. In *IEEE VTC*, volume 5, pages 3147–3151, May-Jun. 2005-Spring.
- [11] F. Saadane, A. Menouni, T. Knopp, and T. Aboutajdine. Empirical eigenanalysis of indoor UWB propagation channels. In *IEEE Globecom*, volume 5, pages 3215–3219, Nov.-Dec. 2004.
- [12] Akbar Sayeed, Vasanthan Ragahvan, and Jayesh Kotecha. Capacity of space-time wireless channels: A physical perspective. In *IEEE Information Theory Workshop*, Oct. 2004.
- [13] I. Emre Telatar and David N. C. Tse. Capacity and mutual information of wideband multipath fading channels. *IEEE Transactions on Information Theory*, 46(4):1384–1400, Jul. 2000.
- [14] M. Z. Win and R. A. Scholtz. Impulse radio: how it works. *IEEE Communications Letters*, 2(2):36–38, Feb 1998.